



# Convolutional dictionary learning for blind deconvolution of optical coherence tomography images

JUNZHE WANG,<sup>1</sup>  BRENDT WOHLBERG,<sup>2</sup> AND R. B. A. ADAMSON<sup>1,3,\*</sup>

<sup>1</sup>*School of Biomedical Engineering, Dalhousie University, Halifax, NS B3H 4R2, Canada*

<sup>2</sup>*Theoretical Division, Los Alamos National Laboratory, Los Alamos, NM 87545, USA*

<sup>3</sup>*Electrical & Computer Engineering Department, Dalhousie University, Halifax, NS B3H 4R2, Canada*

\*[rob.adamson@dal.ca](mailto:rob.adamson@dal.ca)

**Abstract:** In this study, we demonstrate a sparsity-regularized, complex, blind deconvolution method for removing sidelobe artefacts and stochastic noise from optical coherence tomography (OCT) images. Our method estimates the complex scattering amplitude of tissue on a line-by-line basis by estimating and deconvolving the complex, one-dimensional axial point spread function (PSF) from measured OCT A-line data. We also present a strategy for employing a sparsity weighting mask to mitigate the loss of speckle brightness within tissue-containing regions caused by the sparse deconvolution. Qualitative and quantitative analyses show that this approach suppresses sidelobe artefacts and background noise better than traditional spectral reshaping techniques, with negligible loss of tissue structure. The technique is particularly useful for emerging OCT applications where OCT images contain strong specular reflections at air-tissue boundaries that create large sidelobe artefacts.

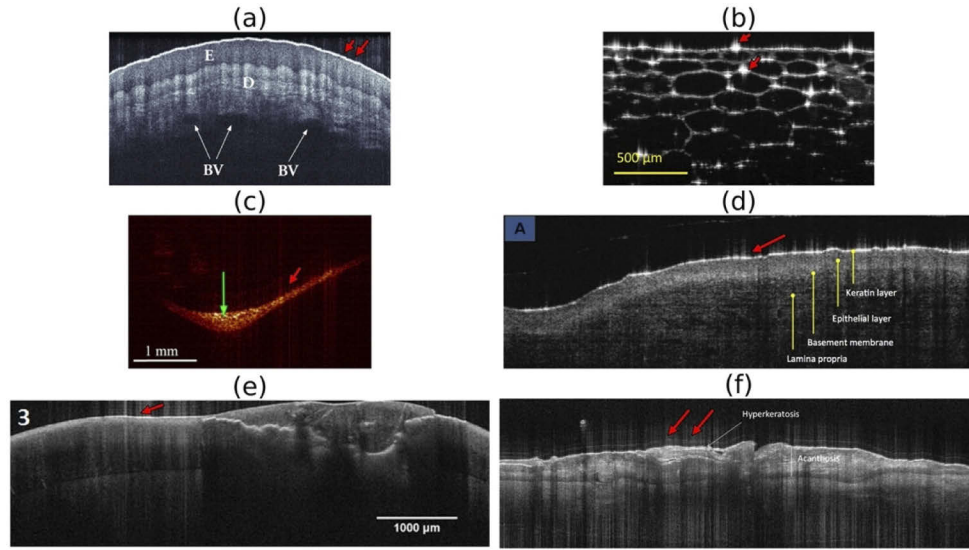
© 2022 Optica Publishing Group under the terms of the [Optica Open Access Publishing Agreement](#)

## 1. Introduction

Optical coherence tomography (OCT) is an imaging modality that captures depth-resolved images with micron resolution by measuring optical interference between light back-reflected from tissue and light in a reference arm [1]. OCT images are often affected by sidelobe artefacts caused by spectral non-uniformity in the OCT light source and/or spectral non-linearity [2,3]. Sidelobe artefacts generally appear as lines extending out axially from bright reflectors and can be seen contaminating many OCT images in the published literature [4–6]. Six examples of published images containing prominent sidelobe artefacts are presented in Fig. 1.

Because they are associated with bright reflectors, sidelobe artefacts are particularly troublesome in emerging OCT applications such as dermatology [7], gastroenterology [7], lung imaging [8], imaging of the oral [9] and respiratory tract mucosa [10], dental imaging [11], and imaging of the middle ear [12,13] where highly reflective air-tissue boundaries form part of the image. In this study we present a novel technique for suppressing these artefacts using sparsity-regularized blind deconvolution.

Most modern OCT systems use Fourier-domain OCT (FD-OCT) techniques [20] in which a spectral interferogram signal is collected by performing a spectrally resolved measurement of the interference intensity between light reflected from the sample and light from a fixed reference arm. Raster-scanned image data is collected on a line-by-line basis with one-dimensional (1D), depth-resolved image lines called A-lines stitched together to form cross-sectional two-dimensional (2D) or volumetric three-dimensional (3D) images.



**Fig. 1.** Published image examples containing prominent sidelobe artefacts (indicated by red arrows). (a) OCT cross-sectional image of a human finger, taken with Thorlabs Inc. Telesto Series OCT SD-system (Image courtesy of Thorlabs Inc.) [14]. (b) OCT cross-sectional image of an onion slice, taken with Wasatch Photonics Inc SD-system with a Cobra 800 spectrometer (Image courtesy of Wasatch Photonics Inc.) [15]. (c) OCT cross-sectional image of human middle ear, taken with a SS-OCT system [16]. (d) OCT cross-sectional image of oral squamous cell carcinoma, taken with Michelson Diagnostics EX1301 OCT Microscope V1.0 (Reprint permission granted by Elsevier) [17]. (e) OCT cross-sectional image of human tooth, taken with VivoSight multiple-beam SS-OCT (Published by SPIE, available under Creative Commons Attribution 4.0) [18]. (f) OCT cross-sectional of human skin with an actinic keratosis with hyperkeratosis, taken with VivoSight multiple-beam SS-OCT (Reprint permission granted by SPIE) [19].

Neglecting constant and autocorrelation terms, the measured OCT spectral interferogram signal  $I(k)$  is given by [21,22]

$$I(k) = G(k) \left( \int_{-\infty}^{\infty} \alpha(z) \cos(2knz) dz \right) \quad (1)$$

where  $G(k)$  is the optical spectrum of the source,  $\alpha(z)$  is the complex scattering amplitude of the sample as a function of depth  $z$ ,  $k = 2\pi/\lambda$  is the wavenumber at the source center wavelength  $\lambda$  and  $n$  is the refractive index [21]. The complex A-line  $s(z)$  is usually calculated via inverse discrete Fourier transformation (DFT),

$$s(z) = F^{-1}\{I(k)\} = F^{-1}\{G(k)\} * \alpha(z) \quad (2)$$

where the  $*$  operator denotes convolution. We can interpret the factor  $F^{-1}\{G(k)\}$  as an axial point spread function (PSF)  $d(z)$  that blurs out the response from a point scattering target in the A-line signal and potentially introduces sidelobe artefacts into the image. Since source spectral imperfections can affect both magnitude and phase,  $d(z)$  is generally a complex function of depth given by

$$d(z) = F^{-1}\{G(k)\} \quad (3)$$

so that the A-line signal  $s(z)$  has a convolutional structure given by [1]

$$s(z) = F^{-1}\{I(k)\} = d(z) * \alpha(z) \quad (4)$$

A primary goal of OCT image processing is to obtain an accurate estimate of the scattering amplitudes  $\alpha(z)$  from the measured  $s(z)$ . Conventionally,  $s(z)$  is taken as the estimator of  $\alpha(z)$  which is an effective approximation as long as the PSF is spatially localized (i.e., it approximates a delta function). However, in practice, OCT axial PSFs are not perfectly localized. Rather PSFs exhibit sidelobe artefacts caused by non-uniformity of the light source spectrum and/or imperfect linearity in spatial frequency of the swept light source in swept source OCT or detector system in spectral domain OCT [3]. In well-designed OCT systems sidelobes can be 40 dB [1] or more lower than the main PSF peak, but when there is a bright reflector on a given line, the PSF sidelobes can rise above the noise floor and appear as artefacts extending from the reflector in the depth direction [1,23,24]. These sidelobes artefacts can obscure structures of diagnostic interest, limit contrast and degrade image quality.

Traditionally, sidelobe artefacts have been addressed primarily using spectral reshaping techniques. Spectral windowing is the simplest and most widespread such technique [1], although more sophisticated windowing approaches have also been used [25]. However, spectral reshaping techniques introduce trade-offs between axial resolution and sidelobe suppression [1]. Furthermore, as linear techniques they are limited in the degree of sidelobe suppression that they can achieve and have difficulty accurately removing artefacts from sharp image features such as tissue boundaries. Because these sidelobe artefacts arise from convolution of the scattering amplitude with the axial PSF, if the axial PSF of an OCT system is known, deconvolution techniques such as Lucy-Richardson [26,27], CLEAN [28], or gradual iterative subtraction (GIS) [4] can be used to suppress the artefacts by deconvolving the axial PSF from the measured A-lines, resulting in improved image quality. In addition to suppressing sidelobes, deconvolution can also be used to achieve improvements to resolution and noise [26].

However, these conventional deconvolution methods require *a priori* knowledge of the PSF from either numerical modeling [29] or experimental measurements [26]. Such methods therefore have difficulty in addressing sidelobe artefacts when the PSF varies over time due to thermal and mechanical drift [6] and they require system-specific calibration, adding complexity and cost to the system design. Depending on the implementation, measuring the PSF may also entail interrupting imaging, causing a decrease in frame rate.

A preferred approach, which we demonstrate for the first time in this study, is to estimate and then deconvolve the axial PSF using only data collected as part of normal imaging, performing so-called blind deconvolution in which the axial PSF is determined from imaging data rather than from a separate measurement. While blind deconvolution techniques are common in image processing [30] and have been previously been applied to improving lateral resolution in OCT [31,32] to the best of our knowledge this is the first application of blind deconvolution of the complex PSF from OCT A-lines to address the problem of axial sidelobe artefacts. Our sparsity-regularized approach to blind deconvolution of OCT A-lines is model-free and requires no modification to hardware or image sequencing. The method can be applied to already-collected OCT data, so long as the complex A-line data or spectrogram data was retained.

## 2. Convolutional sparse coding

If a signal is known to have a convolutional structure – i.e., to consist of a PSF convolved with an unknown vector – and the unknown vector is known to be sparse, i.e., to have a significant fraction of its elements with negligibly small amplitudes, sparse representation methods can be applied to the problem of estimating the unknown vector and the PSF from measured signals [33]. While the sparsity of OCT images is application-specific, many OCT tissue images contain a large fraction of non-reflective regions that would have zero intensity in an ideal image. This makes OCT images good candidates for applying techniques that attempt to reconstruct a maximally sparse estimate of the image consistent with the data.

Over the last 20 years, sparse representation techniques that employ optimization over an  $\ell_1$  norm to find the sparsest representation of a vector within an overcomplete dictionary [34] have been widely applied across many fields of signal processing for denoising [35], compressed sensing [36], interpolation and unsupervised machine learning [37]. The general framework for sparse optimization can be applied to the problem of deconvolution in which a measured signal is assumed to be formed as the convolution of a PSF, typically a short pattern that is repeated throughout the signal, and a sparse vector containing information about where in the signal the pattern occurs and its local amplitude [38]. In cases where the PSF is not known *a priori*, dictionary learning approaches can be used to estimate the PSF from examples of the measured data [38] which turns the sparse estimation into a blind deconvolution method. In recent years, such approaches have been extensively developed in the fields of image processing [39–41], analysis of neural signals [42], and in the analysis of music [43,44].

Sparse representations have previously been used in the processing of OCT images. Compressed sensing techniques that use sparse representations have been applied to reduce the number of spectral or spatial samples needed to form OCT images [45,46]. Sparsity-based denoising [47] and image compression [48] techniques have also been developed for processing of B-mode OCT images. Sparse estimation techniques have also been applied to A-line reconstruction [49,50] to improve axial resolution. However, to our knowledge, sparsity-based processing approaches have not previously been applied to the deconvolution of A-line OCT data. A-line deconvolution requires the extension of sparse representation methods originally developed for real-valued signals to allow for both a complex PSF and a complex sparse vector. It also requires methods for ensuring that sparse representations correctly capture tissue speckle, which is an inherently non-sparse and essential feature of OCT imaging data.

In this study we present a novel blind deconvolution method for complex OCT A-lines based on a sparse representation that is tolerant of tissue speckle. We show that our method is effective in simultaneously suppressing sidelobe artefacts and noise from OCT images while preserving tissue speckle.

### 2.1. Convolutional sparse coding for A-line deconvolution

We approach the deconvolution of OCT A-lines by representing A-line estimation using the convolutional sparse coding (CSC) problem, a well-studied problem in sparse estimation theory [51]. The CSC problem is a minimization problem over the vector  $\mathbf{x}(z)$  expressed as

$$\arg\min_{\mathbf{x}} \left\{ \frac{1}{2} \|\mathbf{s}(z) - \mathbf{d}(z) * \mathbf{x}(z)\|_2^2 + \lambda \|\mathbf{x}(z)\|_1 \right\} \quad (5)$$

where the cost function in [Eq. (5)] includes an  $\ell_2$  term expressing the fidelity between an input A-line signal  $\mathbf{s}(z)$  and the convolution between an axial PSF  $\mathbf{d}(z)$  and an estimate of the scattering amplitude  $\mathbf{x}(z)$  and an  $\ell_1$  regularization term that enforces the sparsity of  $\mathbf{x}(z)$ . A regularization parameter  $\lambda$  determines the relative weighting between the fidelity and sparsity terms. Larger values of  $\lambda$  will promote the sparsity of the estimate  $\mathbf{x}(z)$  at the expense of fidelity while smaller values will favour fidelity between  $\mathbf{s}(z)$  and  $\mathbf{d}(z) * \mathbf{x}(z)$  at the expense of sparsity. An acceptable value of  $\lambda$  generally needs to be determined heuristically on the class of signals being sparsely estimated [52]. [Equation (5)] is a simplified form of the more general CSC problem which involves a sum over multiple convolutional terms [51].

Because the CSC problem is designed to find a sparse estimate of each OCT A-line, its success depends on the scattering amplitude  $\alpha(z)$  of the tissue actually being sparse. In many applications, OCT A-lines are macroscopically sparse [53], containing relatively few brightly reflecting structures and large regions with no reflectors. However, within tissue regions, OCT images exhibit speckle, a temporally static, spatially random texture that is not microscopically sparse [46,54]. When sparsely regularized optimization methods are applied to speckle-containing

images, the optimization tends to cause loss of brightness when low-intensity points within a speckle-containing region are set to zero intensity to increase the sparsity of the estimate  $\mathbf{x}(z)$ . Within the CSC framework, this effect can be mitigated by using a spatially varying weighting of the  $\ell_1$  term in the cost function whereby a larger weight is applied to the  $\ell_1$  term in regions that are free of structure and a smaller weight is applied in regions containing structure. This weighted form of the CSC problem is given by [39]

$$\operatorname{argmin}_{\mathbf{x}} \left\{ \frac{1}{2} \|\mathbf{s}(z) - \mathbf{d}(z) * \mathbf{x}(z)\|_2^2 + \lambda \|\mathbf{W}(z)\mathbf{x}(z)\|_1 \right\} \quad (6)$$

where  $\mathbf{W}(z)$  is a depth-dependent weighting factor applied to the  $\ell_1$  term.

## 2.2. Convolutional dictionary learning for axial point spread function estimation

However, applying CSC-based optimization as a sparsity-regularized deconvolution method still requires that the PSF  $\mathbf{d}(z)$  be known *a priori* and provided as an input to the optimization process. This means that as we have formulated it, CSC is a deconvolution method rather than a blind deconvolution method. To develop CSC into a blind deconvolution method, the CSC problem can be extended to also include optimization over the PSF  $\mathbf{d}(z)$ . This enhancement is called convolutional dictionary learning (CDL) [55]. In CDL, the PSF is determined through optimization over a set of  $K$  OCT A-lines. The CDL problem can be written as

$$\begin{aligned} \operatorname{argmin}_{\mathbf{d}, \mathbf{x}_k} \left\{ \frac{1}{2} \sum_{k=1}^K \|\mathbf{s}_k(z) - \mathbf{d}(z) * \mathbf{x}_k(z)\|_2^2 + \lambda \sum_{k=1}^K \|\mathbf{x}_k(z)\|_1 \right\} \\ \text{s.t. } \|\mathbf{d}(z)\|_2 = 1 \end{aligned} \quad (7)$$

where the minimization runs over both a PSF  $\mathbf{d}(z)$  and a set of  $K$  sparse A-line estimates  $\mathbf{x}_k(z)$  such that the average error between the sparse representations  $\mathbf{d}(z) * \mathbf{x}_k(z)$  and a set of  $K$  measured A-lines  $\mathbf{s}_k(z)$  is minimized. When applied to OCT A-lines collected over a short enough period that the PSF remains stable, the  $\mathbf{s}_k(z)$  can consist of A-lines at different lateral locations within an image in which case  $k$  indexes the lateral location.

The solution to this minimization problem consists of a set of sparse vectors  $\mathbf{x}_k(z)$  that are estimates of the scattering amplitude  $\alpha_k(z)$  at each of  $K$  lateral locations and an estimate of the axial PSF  $\mathbf{d}(z)$ , assumed to be independent of lateral location [24]. In determining  $\mathbf{d}(z)$ , the optimization identifies axially shift-invariant patterns that are common across and within the measured signals  $\mathbf{s}_k(z)$  while ignoring features like background noise that are spatially random across and within A-lines. Used in this way, the CDL problem of [Eq. (7)] amounts to a sparsity-regularized blind axial deconvolution of a complex 2D OCT image.

The CDL problem can be efficiently solved by alternately optimizing for the  $\mathbf{x}_k(z)$  with  $\mathbf{d}(z)$  fixed in a *sparse coding step* and optimizing for  $\mathbf{d}(z)$  with  $\mathbf{x}_k(z)$  fixed in a *dictionary update step* [56]. For efficient solution of both the CSC and the CDL problems, one can use (among other options [55,57]) the alternating direction of multipliers method (ADMM) [51].

We describe the ADMM method as applied to the CDL problem below. ADMM solves the CSC problem in a similar way, only with a fixed, known  $\mathbf{d}(z)$  and a single measured A-line  $\mathbf{s}(z)$  and a single sparse vector  $\mathbf{x}(z)$  in lieu of the ensembles of  $K$  measured A-lines  $\mathbf{s}_k(z)$  and sparse vectors  $\mathbf{x}_k(z)$  used in CDL.

## 2.3. Convolution dictionary learning and convolutional sparse coding in the complex domain

For complex-valued  $\mathbf{x}_k(z)$  and  $\mathbf{d}(z)$ , the sparse coding step in conventional CDL ADMM solvers designed for real-valued signals must be modified to work with complex vectors as originally



proposed in [58]. The cost function of [Eq. (7)] can be readily generalized to complex  $\mathbf{x}_k(z)$  and  $\mathbf{d}(z)$  by defining the  $\ell_1$  and  $\ell_2$  norms on complex arguments as

$$\|\mathbf{x}_k(z)\|_1 = \sum_{i=1}^N |\mathbf{x}_k(z_i)| \quad (8)$$

and

$$\|\mathbf{s}_k(z) - \mathbf{d}(z) * \mathbf{x}_k(z)\|_2^2 = \sum_{i=1}^N |\mathbf{s}_k(z_i) - \mathbf{d}(z_i) * \mathbf{x}_k(z_i)|^2 \quad (9)$$

where  $N$  is the number of pixels in the axial direction,  $|\cdot|$  denotes the magnitude of a complex number and  $\mathbf{s}_k(z)$ ,  $\mathbf{d}(z)$  and  $\mathbf{x}_k(z)$  are all complex vectors.

In the ADMM method, the sparse coding step is solved by converting [Eq. (7)] to matrix form

$$\operatorname{argmin}_X \left\{ \frac{1}{2} \|\mathbf{D}\mathbf{X} - \mathbf{S}\|_2^2 + \lambda \|\mathbf{X}\|_1 \right\} \quad (10)$$

where  $\mathbf{D} \in \mathbb{C}^{N \times N}$  is the matrix form of the convolution operator with  $\mathbf{d}(z)$  such that  $\mathbf{d}(z) * \mathbf{x}_k(z) = \mathbf{D}\mathbf{x}_k$ ,  $\mathbf{S} = [\mathbf{s}_1(z), \dots, \mathbf{s}_K(z)] \in \mathbb{C}^{N \times K}$  is a matrix formed from the measured A-lines  $\mathbf{s}_k(z)$ , and  $\mathbf{X} = [\mathbf{x}_1(z), \dots, \mathbf{x}_K(z)] \in \mathbb{C}^{N \times K}$  is a matrix formed from the sparse vectors  $\mathbf{x}_k(z)$ .  $\mathbf{D}\mathbf{X}$  is then the sparse estimate of  $\mathbf{S}$ .

The problem is then transformed into a constrained optimization problem which is solved by minimization of an augmented Lagrangian function containing a dual variable  $\mathbf{U} \in \mathbb{C}^{N \times K}$  and an auxiliary variable  $\mathbf{Y} \in \mathbb{C}^{N \times K}$ . A real, positive parameter  $\rho$  governs the rate of convergence. In this study  $\rho$  is selected using the adaptive method described in [38].

$$L_\rho(\mathbf{X}, \mathbf{Y}, \mathbf{U}) = \frac{1}{2} \|\mathbf{D}\mathbf{X} - \mathbf{S}\|_2^2 + \lambda \|\mathbf{Y}\|_1 + \frac{\rho}{2} \|\mathbf{X} - \mathbf{Y} + \mathbf{U}\|_2^2 \quad (11)$$

The minimization process is performed according to the following steps:

2.3.1. Find  $\mathbf{X}$  that minimizes  $L_\rho$  with  $\mathbf{Y}$  held constant

This problem can be formulated as

$$\operatorname{argmin}_X \left\{ \frac{1}{2} \|\mathbf{D}\mathbf{X} - \mathbf{S}\|_2^2 + \frac{\rho}{2} \|\mathbf{X} - \mathbf{Y} + \mathbf{U}\|_2^2 \right\} \quad (12)$$

This minimization can be solved by setting the derivative of  $L_\rho$  to zero, leading to a linear equation system

$$(\mathbf{D}^H \mathbf{D} + \rho \mathbf{I})\mathbf{X} = \mathbf{D}^H \mathbf{S} + \rho(\mathbf{Y} - \mathbf{U}) \quad (13)$$

where the  $H$  superscript denotes the Hermitian transpose. The equation system can be solved in the frequency domain by taking the discrete Fourier transform (DFT) of both sides

$$(\hat{\mathbf{D}}^H \hat{\mathbf{D}} + \rho \mathbf{I})\hat{\mathbf{X}} = \hat{\mathbf{D}}^H \hat{\mathbf{S}} + \rho(\hat{\mathbf{Y}} - \hat{\mathbf{U}}) \quad (14)$$

where  $\hat{\mathbf{A}}$  is the DFT of  $\mathbf{A}$ . The solution can then be obtained by exploiting the Sherman-Morrison formula [38].

2.3.2. Find  $Y$  that minimizes  $L_p$  with  $X$  held constant

This problem can be written as

$$\operatorname{argmin}_Y \left\{ \lambda \|Y\|_1 + \frac{\rho}{2} \|X - Y + U\|^2 \right\} \quad (15)$$

This is the step that must be modified for complex variables. When  $X$ ,  $Y$  and  $U$  are real-valued, the problem has a closed-form solution

$$Y = S_{\frac{\lambda}{\rho}}(X + U) \quad (16)$$

where  $S(\cdot)$  is the soft thresholding operator defined for real-valued  $A$  as

$$S_{\gamma}(A) = \operatorname{sign}(A) \odot \max(0, |A| - \gamma) \quad (17)$$

where  $\odot$  denotes element-wise multiplication and the  $\operatorname{sign}(\cdot)$  and  $\max(\cdot)$  functions are applied elementwise.

However, the  $\operatorname{sign}(\cdot)$  operator is not defined on complex numbers. Following [58], for  $A \in \mathbb{C}$  we define a complex soft thresholding function

$$CS_{\gamma}(A) = \frac{A}{|A|} \odot \max(0, |A| - \gamma) \quad (18)$$

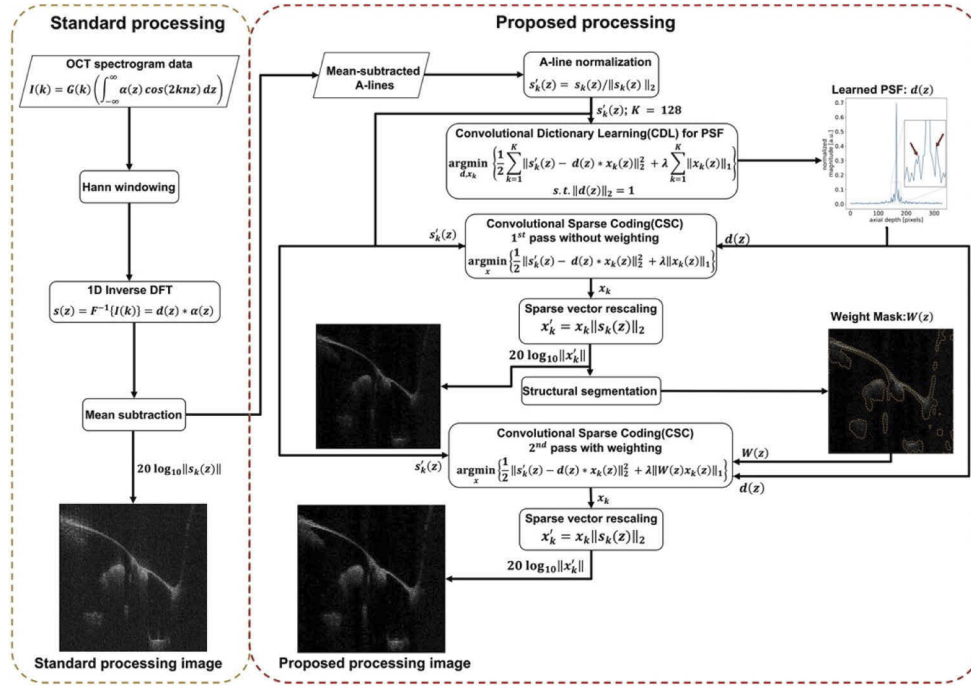
which is equivalent to [Eq. (16)] when  $A$  is real.  $CS_{\gamma}(A)$  preserves the phase of the complex vector for  $|A| > \gamma$  and collapses vectors with  $|A| \leq \gamma$  to the origin on the complex plane. This function provides a closed form solution to [Eq. (15)] for complex vectors.

The dictionary update step in the ADMM CDL solution is unaffected by whether  $d(z)$ ,  $x(z)$  and  $s(z)$  are complex, and therefore does not require modification. A full derivation of the ADMM solution to the CDL problem with complex vectors using the complex soft thresholding function can be found in [58].

In this study, we applied the CDL framework to perform blind deconvolution of complex OCT A-lines with the goal of suppressing PSF sidelobe artefacts and noise while preserving structural fidelity. All CDL and CSC processing used the open-source, Python-based SPORCO toolkit [59] which, for this study, was extended to support complex-valued sparse estimation using complex soft thresholding. SPORCO supports several solvers, but all results in this study use its CDL ADMM solver or its CSC ADMM solver. All source code and example data for the methods applied in this study can be accessed from our public GitHub repository [60].

### 3. Methods

OCT spectrogram data was collected using a previously described [12] OCT system designed around a 1550nm aperiodic swept laser source (Insight Photonics SLE-101) with a sweep bandwidth of 40nm, axial resolution in air of 40 $\mu$ m, lateral resolution in air at the focus of 35 $\mu$ m and a nominal sweep rate of 100 kHz. The beam was scanned laterally using a 2D MEMS mirror (Mirrorcle A8L2.2). The system was designed for imaging of the middle ear and by design it has lower axial and lateral resolution than typical ophthalmic OCT systems. In each image dataset, we collected 10,240 A-lines over 102.4 ms and sampled every 20th line to form a B-mode dataset with a lateral width of 512 pixels. We applied a standard set of A-line preprocessing steps to the spectrogram data [1], which we will refer to as *standard processing* in this study. Standard processing consisted of windowing with a Hann window to compensate for the source spectrum non-uniformity, inverse discrete Fourier transformation (DFT) and fixed pattern noise removal by mean subtraction [61]. The two processing methods are shown in flowchart form in Fig. 2. In



**Fig. 2.** Standard and proposed OCT A-line processing methods. The proposed method uses convolutional dictionary learning for axial PSF estimation and convolutional sparse coding for sparse estimation of the tissue structure.

what follows, the reference images have only had these standard processing steps applied while the deconvolved images have had these steps applied before applying the proposed processing.

Because the  $\ell_1$  and  $\ell_2$  terms in the cost functions of [Eq. (5)] and [Eq. (7)] scale differently with  $s_k$ ,  $s_k$  was normalized by its  $\ell_2$  norm on each line prior to performing the sparse optimization so that the same value of  $\lambda$  could be used across all A-lines. That is, we rescaled  $s_k(z)$  to  $s'_k(z) = s_k(z) / \|s_k(z)\|_2$ . Following convolutional sparse coding, we rescaled  $x_k$  to  $x'_k = x_k \|s_k(z)\|_2$  to restore the sparse vector signal amplitude to the level of the measured signal.

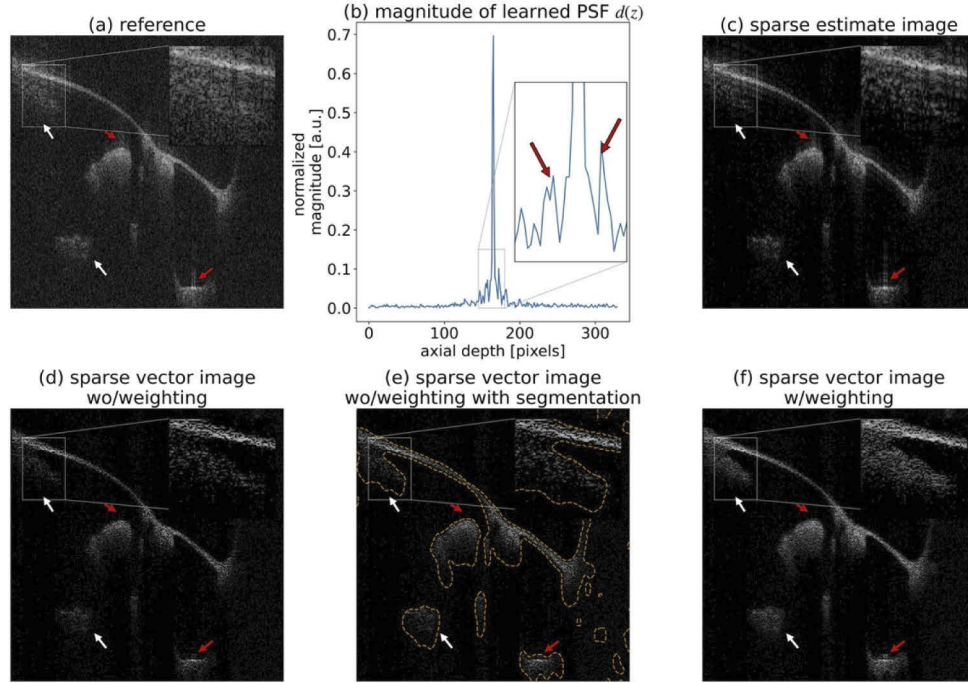
The set of normalized complex A-lines  $s'_k(z)$  form the input to the CDL ADMM solver. In each image, we learned the PSF  $d(z)$  using a subset of lines consisting of every fourth line out of the 512 (so that  $K = 128$ ) using a CDL solver with  $\lambda=0.1$ , and then applied a CSC solver to obtain the sparse vector  $x'_k(z)$  of all 512 lines using the PSF from the CDL solution. Although it is possible to learn the PSF starting from a random initial guess, we found that picking one of the normalized A-lines as the initial guess for the PSF  $d(z)$  accelerated the dictionary learning process while arriving at the same PSF estimate.

In what follows we refer to  $d(z) * x'_k(z)$  as the *sparse estimate* of the A-line and we construct the *sparse estimate image* from  $20 \log_{10} \|d(z) * x'_k(z)\|$  where the lateral location is indexed by  $k$ . We refer to  $x'_k(z)$  as the *sparse vector estimate* of  $\alpha(z)$ , i.e., of the complex tissue scattering amplitudes. We form the *sparse vector image* as  $20 \log_{10} \|x'_k(z)\|$ . The sparse vector image is the final output from our proposed processing method.

Figure 3 shows an example of applying the proposed processing to an OCT image of a human middle ear. All images in this study are displayed on the same intensity scale with a 50dB dynamic range. Figure 3(a) shows a reference image which is affected by sidelobe artefacts proximal to brightly reflecting middle ear structures (indicated by the red arrows) as well as background noise. White arrows show weakly reflecting structures in the image. Figure 3(b)



shows the learned PSF obtained by solving the CDL problem of [Eq. (7)] and highlights the sidelobes (red arrows) that contribute the artifacts in Fig. 3(a) and Fig. 3(c). The insets highlight a region of the image where these artefacts are sufficiently severe to completely obscure an empty region between the tympanic annulus and tympanic bone.



**Fig. 3.** Example of applying the proposed processing to an OCT image of a human middle ear. (a) reference image with only standard OCT processing applied. (b) The magnitude of the point spread function  $d(z)$  learned over a subset of 128 lines within the 512-line image by solving the CDL minimization problem of [Eq. (7)] with  $\lambda = 0.1$ . (c) the sparse estimate image  $d(z) * x'(z)$  obtained by solving the weighted CSC minimization problem of [Eq. (6)]. (d) the sparse vector image obtained from CSC with  $\lambda = 0.05$  and without  $\ell_1$  weighting (i.e., with  $W(z) = 1$  everywhere). (e) sparse vector image from (d) segmented into the tissue-containing regions using Sobel filter-based edge detection (highlighted with dashed orange overlay). (f) sparse vector image obtained by solving [Eq. (6)] with  $W(z) = 0.1$  in the segmented regions of (e), and  $W(z) = 1$  elsewhere with a transition from  $W(z) = 0.1$  to  $W(z) = 1$  taking place over 20 pixels in the regions distal to each segmented region. In all images, red arrows indicate the sidelobe artefacts and white arrows highlight tissue regions of interest.

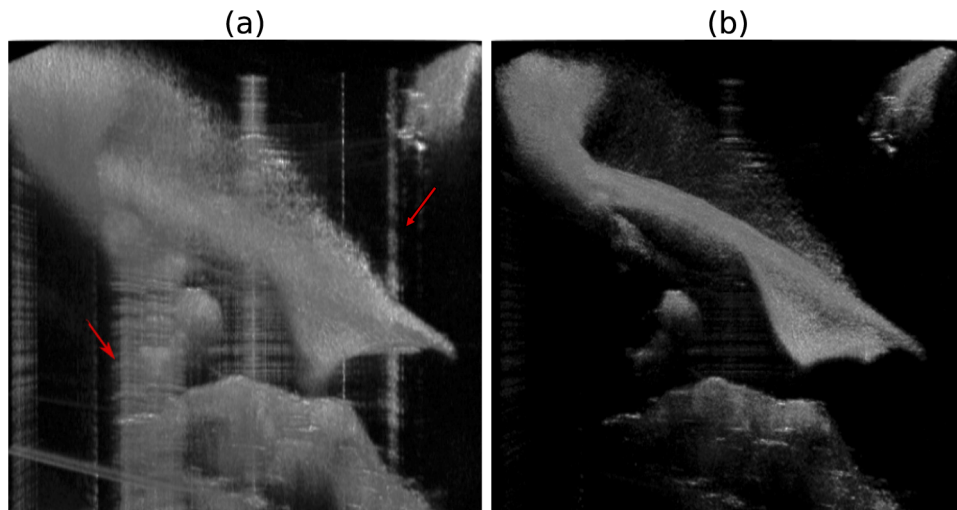
Figure 3(c) shows an image generated from the sparse estimate  $d(z) * x'(z)$  obtained from solving the CSC problem with the learned PSF. The sparse regularization  $\ell_1$  term of [Eq. (5)] drives a substantial reduction in the background noise as  $x'(z)$  is made sparse within the empty regions in the proximal and distal portions of the image. At the same time, the fit  $d(z) * x'(z)$  achieves high fidelity with both the tissue structure and the sidelobe artefacts. Figure 3(d) shows an image constructed from the sparse vectors  $x'(z)$  without applying an  $\ell_1$  weighting factor (i.e., with  $W(z) = 1$ ) and with  $\lambda = 0.05$ . In this image, the sidelobe artefacts are highly suppressed because they are contained entirely in the learned PSF  $d(z)$  and not in the sparse vectors  $x'(z)$ . In the top left quadrant of the image (highlighted in the inset), the empty region that was completely obscured by these sidelobe artefacts in Fig. 3(a) and Fig. 3(c) is revealed in the sparse vector

image of Fig. 3(d). However, in the tissue-containing regions, many pixels have been set to zero intensity by the sparse optimization process to increase the sparsity of the solution, resulting in loss of speckle brightness in the tissue regions.

This is an example of a central problem in applying sparsity-based image processing to images that contain strong speckle, such as OCT or ultrasound images [62]. Because the speckle makes the tissue regions non-sparse, the sparsity-based optimization will tend to set low intensity pixels in the speckle to zero intensity in the sparse vectors  $\mathbf{x}'_k(z)$ , potentially causing a loss of diagnostically useful information. While it would be possible to apply despeckling prior to deconvolution, the non-linear nature of the despeckling process would destroy the convolutional structure of the OCT image resulting in a lower fidelity deconvolution.

An alternate approach to mitigating the loss of speckle brightness is to apply a spatially dependent weighting factor  $W(z)$  to the  $\ell_1$  term in [Eq. (6)] that determines how heavily sparsity will be weighted in different regions of the image. If coarse image segmentation is applied to the image to roughly determine the extent of speckle-containing tissue regions, then this segmentation can be used to weight sparsity less heavily in regions containing tissue as compared to regions that do not contain tissue.

To apply this approach, we first obtained sparse vectors  $\mathbf{x}'_k(z)$  at a fixed  $\lambda = 0.05$ . These sparse vectors were nearly free of sidelobe artefacts due to the deconvolution of the PSF. However, they also suffered from loss of speckle brightness. We then segmented the sparse vector image for tissue structure by performing opening (dilation and erosion), median filtering and Sobel filter-based edge detection [63]. Other common segmentation approaches would serve equally well. The resulting segmentation contours can be seen in orange overlay in Fig. 3(e). We defined a hard step in the weight map  $W(z)$  at the proximal interface of each structure with a value of  $W(z) = 1$  proximal to the interface and  $W(z) = 0.1$  distal to the interface. Because in middle ear OCT images, the distal boundary of structures is not well defined, a tapered weight was used at the distal interface of each structure. From the distal interface obtained from Sobel-filter segmentation, the weight  $W(z)$  was increased linearly over 20 pixels until it equaled 1.



**Fig. 4.** (a) 3D volumetric OCT image of a human middle ear processed using the standard processing of Fig. 2. The red arrows indicate strong sidelobe artefacts arising from bright specular reflections from the tympanic membrane. (b) The same volume processed using the proposed processing on a slice-by-slice basis displayed with the same dynamic range. Sidelobe artefacts are strongly suppressed. See Visualization 1.

Figure 3(f) shows the result of performing a second pass of the CSC optimization (i.e., solving [Eq. (6)]) with this weighting mask (orange overlay seen in Fig. 3(e)). In this final image, the artefacts remain well-suppressed, but without the loss of tissue brightness seen in Fig. 3(d).

Figure 4 shows the results of applying this process to a full volumetric dataset of a human middle ear. Processing the data using standard processing (Fig. 4(a)) results in strong sidelobe artefacts wherever strong, specular reflections occur (red arrows) which obscure structures behind the artefact and clutter the image. More subtly, sidelobe artefacts introduce a haze around all structures at air-tissue boundaries and make thin structures like the tympanic membrane appear thicker than they are. Figure 4(b) shows the results of applying the proposed blind deconvolution method on a frame-by-frame basis within the volume. The sidelobe artefacts are strongly suppressed and previously obscured structures become visible. The tympanic membrane appears thinner because it is no longer being axially smeared out by the PSF. Figure 4 is a still frame of Visualization 1 which shows the standard-processed and deconvolved volumetric datasets being rotated.

#### 4. Performance assessment

A wide range of image quality metrics have been used in evaluating OCT image processing algorithms, particularly despeckling algorithms. The structure similarity index measure (SSIM) [64] measures similarity between a gold standard image and a processed image to determine similarity. It is not an effective measure of the quality of deconvolved images because there is no gold standard image to compare the sparse vector image to. Edge preservation index (EPI) [65] measures the edge distortion and equivalent number of looks (ENL) [66] is a measure of the smoothness of an image within a region of interest (ROI). While useful in evaluating despeckling algorithms, these metrics are of little use in evaluating deconvolution algorithms since deconvolution tends to decrease smoothness by making features better defined, by increasing the apparent resolution of structures [26,28] and by increasing speckle variance [26]. As a result, EPI and ENL scores will nearly always be worse for sparsely deconvolved images, even when these images are perceptually better at capturing the underlying tissue reflectance than the original image.

Given the limitations of these image quality metrics for our use case, we opted to quantitatively assess the performance of our CDL processing pipeline using three simple metrics that have previously been used to evaluate deconvolution algorithms for OCT and/or ultrasound imaging, the signal-to-noise ratio (SNR) [26] the contrast (C) [67] and the generalized contrast-to-noise ratio (gCNR) [68]. In applying these metrics, we compare the image produced using the proposed CDL-based sparse deconvolution and a standard-processed reference image.

Signal-to-noise ratio (SNR) is defined as the ratio of mean intensity in a ROI containing structure to the standard deviation of intensity in a background region free from structure or sidelobe artefacts

$$SNR = 10 \log_{10} \frac{\mu_h}{\sigma_b} \quad (19)$$

where  $\mu_h$  and  $\sigma_b$  represent the mean intensity (in linear units) of a tissue-containing region (e.g., the dashed red rectangle in Fig. 8) and standard deviation of intensity in a background region (e.g., region B in Fig. 8).

Contrast measures the ratio of mean image intensity between two ROIs within the image (e.g., regions  $H_1$  and  $H_2$  in Fig. 8)

$$C = 10 \log_{10} \frac{\mu_{r1}}{\mu_{r2}} \quad (20)$$

gCNR is a bounded metric ranging from 0 to 1 that quantifies distinguishability in terms of the Bayesian probability of correctly assigning a given pixel to one of two defined regions using a threshold-based binary classifier [69]. It is insensitive to changes in speckle statistics and to the

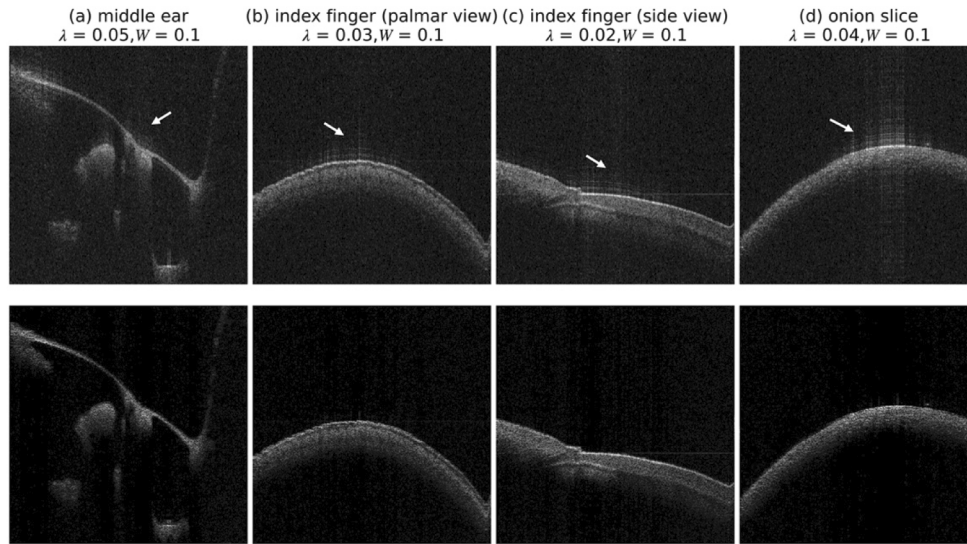
image's dynamic range. To a good approximation, gCNR is a simple function of the overlap integral between the greyscale intensity histograms of two ROIs. A gCNR close to unity indicates that two ROIs can be distinguished with a high confidence and that their intensity histograms are nearly non-overlapping while two regions with gCNR close to zero are nearly indistinguishable and have highly overlapping intensity histograms. Because of its basis in an operational, Bayesian definition of distinguishability, gCNR offers a valid way of assessing change to distinguishability between two regions that is free of any assumptions about the speckle statistics [69]. gCNR is given by

$$gCNR = 1 - \sum_{k=0}^{N-1} \min(r_1(x_k), r_2(x_k)) \quad (21)$$

where  $r_1(x_k)$  and  $r_2(x_k)$  are normalized log-intensity histograms on the  $N = 256$  greyscale levels  $x_k$  of the pixels within the 50 dB display dynamic range in the regions 1 and 2.

## 5. Results and discussions

Figure 5 shows the results of applying the proposed processing to in vivo OCT images of a human middle ear, the palmar aspect of a fingertip, the side aspect of a fingertip, and an onion. All images were processed with a fixed  $\lambda = 0.1$  for the CDL step, followed by the two-pass CSC processing method shown in Fig. 2 with a value of  $\lambda$  selected heuristically for each image to achieve the best perceptual image quality. The top row shows the reference images, processed using the standard processing method. The bottom row shows the results of applying our proposed processing method.



**Fig. 5.** From left to right: OCT images of a middle ear, index finger (palmar view), index finger (side view), and onion slice. The top row shows the reference image with only standard processing applied while the bottom row shows the corresponding sparse vector images obtained from the proposed two-pass sparse processing shown in Fig. 2. The white arrow indicates the sidelobe artifacts caused by the PSF.

The values of the regularization parameter  $\lambda$  and weighting factor  $W(z)$  are given above each image. Qualitatively, in all images, both the sidelobe artefacts (indicated by white arrows) and the background noise level are substantially reduced in the sparse vector images while the contrast of



speckle and tissue boundaries is enhanced. These improvements are achieved without noticeable loss of anatomical structure, even for structures that are weak in the reference images.

For a  $512(w) \times 330(d)$  pixel image, the CDL solver took 750 iterations and 18 seconds to estimate the PSF from 128 A-lines on a 2019 MacBook Pro with a 2.4-GHz Intel i9-9980HK processor and 32-GB of 2400 MHz DDR4 memory. Following dictionary learning of the PSF, the CSC problem was solved using the learned PSF to estimate all 512 image lines on the same computer in a total time of 1.28 seconds. Given the fundamental parallelism of line-by-line based image deconvolution, the computation time of both CDL and CSC could be improved substantially by using parallel processing employing graphics processing unit (GPU) acceleration [59], although that was not explored in this study.

### 5.1. Selection of regularization parameter and $\ell_1$ weighting

Within the proposed processing method, the regularization parameter  $\lambda$  controls the relative weighting of sparsity and fidelity in the CSC and CDL cost functions while the weighting factor  $W(z)$  controls the amount of structure loss in regions identified as structure-containing during segmentation of the sparse vector image from the first CSC pass. Heuristically, we found that  $\lambda$  between 0.01 and 0.05 and  $W = 0.1$  gave good qualitative results across the images we investigated, but we also varied both parameters to investigate the trade-offs from using different values in each image set.

Figure 6 shows the effects of varying  $\lambda$  from  $10^{-5}$  to 0.40 with  $W(z) = 0.1$  for the middle ear image. A tissue structure ROI (red) was selected around the lenticular process of the incus and zoomed in on in the third row. In the case where  $\lambda = 10^{-5}$  the problem of [Eq. (7)] is ill-posed. The sparse estimate overfits the reference image and the sparse vector image is a very poor estimate of the underlying structure. In the range  $\lambda \in [0.01, 0.15]$  the problem is well-posed, and the solution finds a balance between fidelity and sparsity that results in a good estimate of the underlying structure with sidelobe artefacts suppressed. At  $\lambda = 0.01$ , some of the sidelobe artefacts (white arrow) and noise from the reference image are retained in the sparse vector image whereas when  $\lambda = 0.05$  the sidelobe artefact and noise have been completely removed from the empty regions without noticeable tissue loss relative to the reference images. For  $\lambda \geq 0.15$  the algorithm over-weights sparsity causing a loss of tissue speckle brightness (dashed red circle). In one region (green arrow) this loss creates an apparent discontinuity in the tympanic membrane.

Figure 7 demonstrates the effect of applying various values of the weighting factors  $W(z)$  in the segmented tissue regions with  $\lambda$  fixed at 0.05. Increasing  $W(z)$  in the tissue region is equivalent to locally increasing  $\lambda$  and so it also causes progressive loss of signal intensity and a higher fraction of zero intensity pixels in the tissue region indicated by the dashed red circles.

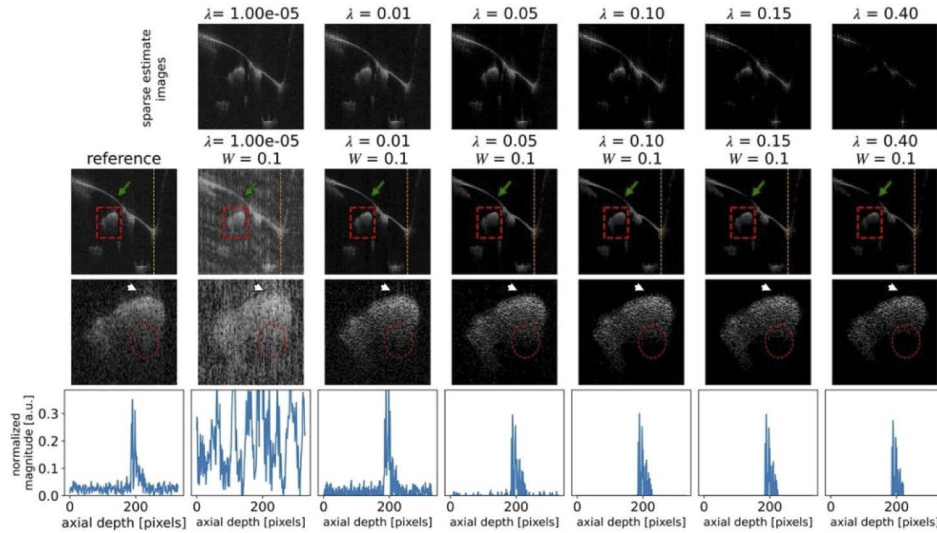
### 5.2. Sparse optimization performance

In Fig. 8, we identified a ROI that contained no structure (cyan, labeled  $B$  for background), a ROI that contained no structure but that was contaminated by sidelobe artefacts from a nearby surface (green, labeled  $A$  for artefact) and two ROIs that contain tissue speckle (red), a brightly reflecting ROI  $H_1$  and a weakly reflecting ROI  $H_2$ .

Our proposed processing improved the SNR of  $H_2$  relative to the background ROI by 2.2 dB and improved the contrast between  $H_2$  and background ROI by 5.8 dB. The processing generally preserved the contrast between tissue regions, with the contrast between  $H_1$  and  $H_2$  changing by only 0.3 dB following processing. gCNR between region  $H_2$  and the background improved modestly from 0.53 to 0.59 and gCNR between region  $H_2$  and the region contaminated by sidelobe artefacts improved substantially from 0.28 to 0.65. However, the sparse processing also caused a substantial decrease in gCNR between  $H_1$  and  $H_2$  from 0.80 to 0.60.

The decrease in gCNR between  $H_1$  and  $H_2$  can be explained almost entirely by an increase in the number of pixels in  $H_1$  with greyscale intensity zero created by the sparse optimization

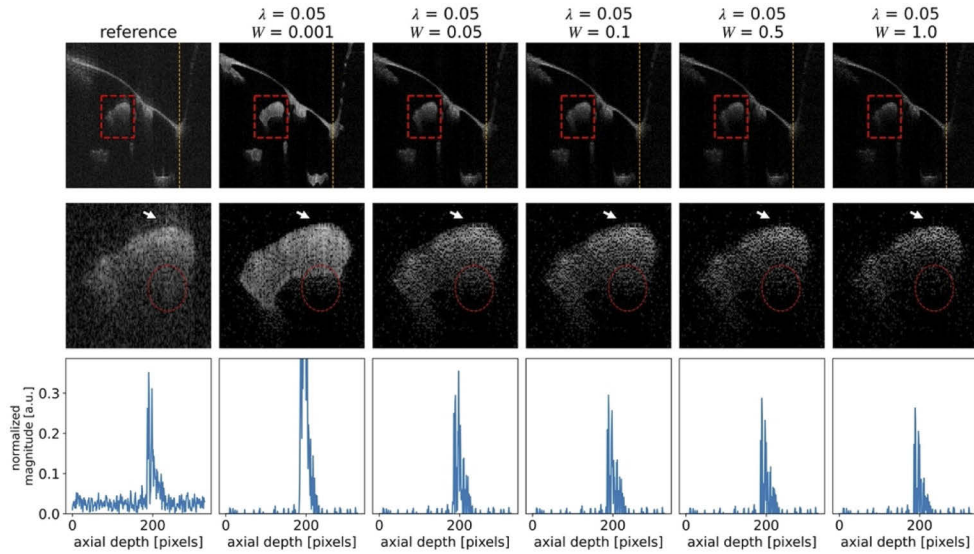




**Fig. 6.** Sparse reconstructions of an OCT middle ear image using the same learned PSF for various values of the regularization parameter  $\lambda$ . The top row shows the sparse estimate images. The second row shows the sparse vector images with the weighting mask applied. The red box highlights a bony tissue ROI (lenticular process of the incus) which is zoomed in the third row. The white arrow indicates the location of sidelobe artefacts appearing in the reference image and suppressed in the sparse images. The dashed red circle highlights a weakly reflecting speckle region at the distal side of the ROI where a loss of structure can be seen as the value of  $\lambda$  increases. The green arrow highlights a region of the tympanic membrane that is completely eroded away in the high  $\lambda$  images. The bottom plots show the A-line magnitude along a selected line in the image (orange line in the second-row images) for each value of  $\lambda$ .

as can be seen in the intensity histograms in the second row of Fig. 8. Because in the sparse representation both  $H_1$  and  $H_2$  have a large fraction of zero-intensity pixels, the histograms have substantial overlap (i.e., orange shading) and the resulting gCNR is low. Because the observed decrease in gCNR is due to the zero-intensity pixels, further processing steps that fill in zero-intensity pixels with non-zero intensity (i.e., despeckling operations) can mitigate the decrease. For example, applying a 2D median filter with a  $3 \times 3$  square kernel to the sparse vector image fills in most of the zero-intensity pixels in the tissue regions and results in a substantial increase in the gCNR between  $H_1$  and  $H_2$  from 0.60 to 0.87 due to the decrement in the number of zero intensity pixels. It therefore appears that a processing sequence consisting of sparse deconvolution followed by despeckling can result in removal of sidelobe artefacts and suppression of background noise without compromising tissue distinguishability.

To study how the gCNR between image regions is affected by the regularization parameter  $\lambda$ , we applied  $\lambda$  values from  $10^{-4}$  to 1 to the index finger (palmar view) image of Fig. 8 and calculated the gCNR between regions at each  $\lambda$ . For  $\lambda < 10^{-4}$  the regularization provided by the sparsity term in [Eq. (7)] is inadequate and the problem becomes ill-posed leading to solutions like the one seen in the second column of Fig. 6. Starting around  $\lambda = 10^{-4}$ , the problem becomes well-posed and the sparse vector image exhibits comparable or improved distinguishability compared to the reference image between the tissue regions, and the background and artefact regions as measured by  $gCNR_{H_1/B}$ ,  $gCNR_{H_2/B}$ ,  $gCNR_{H_1/A}$  and  $gCNR_{H_2/A}$ . As  $\lambda$  increases, all regions become sparser with a greater fraction of their pixels in the zero-intensity histogram bin. The effect of this change on gCNR depends on the regions involved.  $gCNR_{H_2/A}$  is improved while

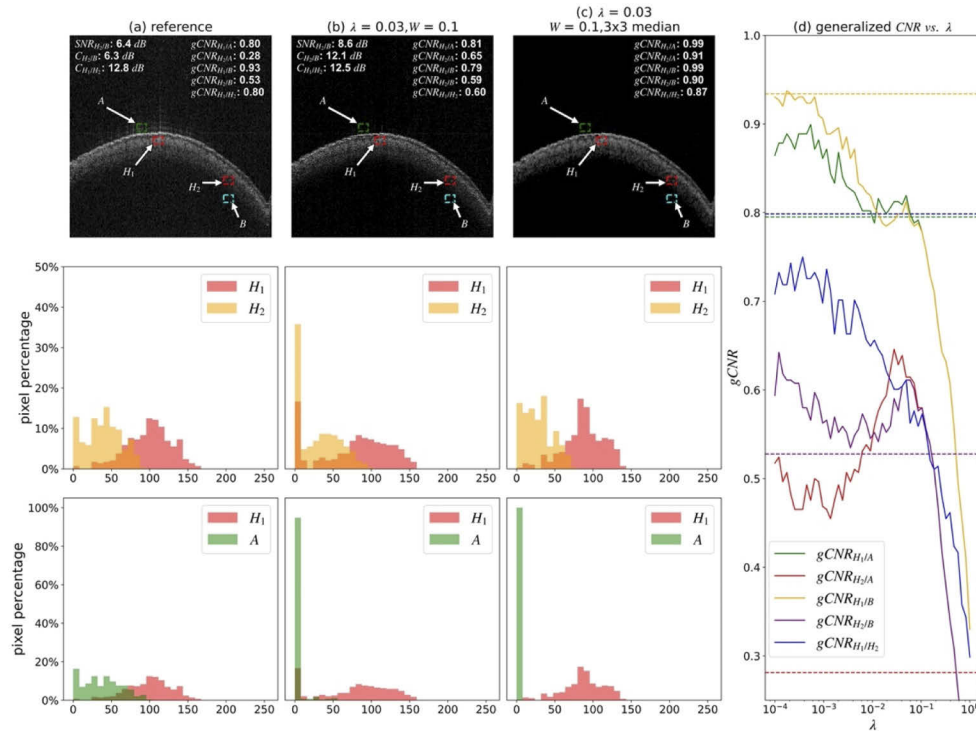


**Fig. 7.** Sparse reconstructions of the OCT middle ear image of Fig. 6 for various values of the weighting factor  $W$  at fixed  $\lambda = 0.05$ . Top row: reference and sparse vector images. A ROI at the lenticular process of the incus is highlighted in the red box. Middle row: zoomed in image of the ROI from the top images. Bottom row: sparse vector A-line estimate along the orange line in the top row images.

$gCNR_{H_1/A}$  and  $gCNR_{H_1/H_2}$  decrease due to the overlap of the zero-intensity bin.  $gCNR_{H_2/B}$  and  $gCNR_{H_2/A}$  both reach their peak at around  $\lambda = 0.03$ . At around  $\lambda = 0.08$ , all  $gCNR$  values begin to roll off as all regions begin to contain a large fraction of zero-intensity pixels, making them less distinguishable from each other by the  $gCNR$  metric. Figure 8(d) supports the heuristic selection of  $\lambda$  in the range 0.01 to 0.05 since, according to the plot, this range offers good artefact suppression while keeping the loss of distinguishability between the two tissue regions due to zero-intensity pixels to an acceptable level.

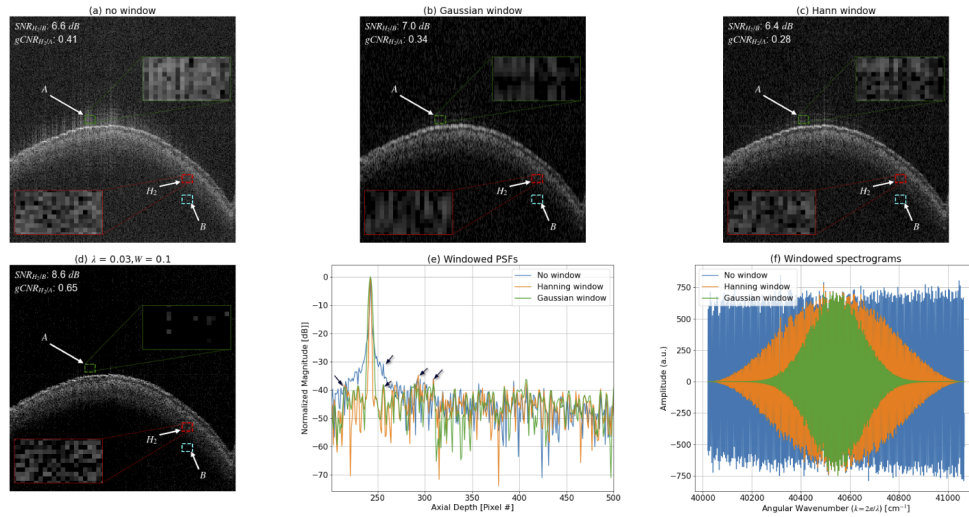
Traditionally, the main method of addressing sidelobe artefacts in the absence of an independent measurement of the PSF has been through the application of spectral windowing [25] prior to inverse discrete Fourier transformation of the spectral interferogram. While simple to implement and computationally efficient, spectral windowing incurs a trade-off between axial resolution and sidelobe suppression that limits its effectiveness [1].

We compared our proposed method to spectral windowing with two frequently used window functions, the Gaussian window, and the Hann window and to processing without windowing. Figure 9(a)-(c) show images produced with, respectively, no spectral windowing, a Gaussian window with the same length as the raw spectrogram and a standard deviation of 10% of the spectrogram width and a Hann window. The Gaussian window was made narrow to highlight the degradation in axial resolution from windowing. While windowing does reduce the severity of sidelobe artefacts relative to no windowing close to the main lobe, the proposed method, shown in Fig. 9(d), significantly outperforms windowing, particularly for artefacts lying 20 pixels or more from the main lobe. The inability of traditional windowing to suppress these artefacts is also highlighted in Fig. 9(e) which shows the directly measured axial PSF from a mirror reflector, calculated with various windows applied. While windowing successfully reduces the sidelobes directly adjacent to the main lobes, other lobes, presumably arising from source spectrum imperfections and/or sweep non-linearity are not suppressed by windowing. Figure 9(f) shows the spectrogram corresponding to the three PSFs in Fig. 9(e).



**Fig. 8.** Index finger (palmar aspect) images with  $\lambda = 0.03$  and  $W = 0.1$ . The performance of the sparse processing algorithm is assessed within three classes of region of interest (ROIs): background region  $B$  (cyan) where there is no anatomical structure, region  $A$  (green) containing sidelobe artefacts, a brightly reflective tissue speckle region  $H_1$  (red) and a weakly reflective tissue speckle region  $H_2$  (red). OCT 2D images are shown in the top row and corresponding intensity histograms in the bottom two rows. (a) reference image (b) image after proposed processing (c) image after proposed processing and median filtering with a  $3 \times 3$  kernel (d) plot of gCNR between the ROIs as a function of the regularization parameter  $\lambda$ . The dashed lines show the gCNR values calculated for the reference image and the solid lines show gCNR for the sparse vector image as a function of  $\lambda$ , with  $W$  held fixed at 0.1.

The superior artefact suppression of the proposed method allows it to achieve higher gCNR between the weak tissue ROI  $H_2$  and sidelobe artefact ROI  $A$ , and a larger SNR improvement between the weak tissue ROI  $H_2$  and background region  $B$  than the windowing methods without noticeable degradation in axial resolution relative to the no-window case. The counterintuitive drop in  $gCNR_{H_2/A}$  between Fig. 9(a) and Fig. 9(b)-(c) can be attributed to the artefact region being brighter than  $H_2$  in Fig. 9(a) which makes the two regions more distinguishable than in Fig. 9(b)-(c) where they have similar intensities. Figure 9(b) shows the degradation in axial resolution that comes from the use of a narrow Gaussian window. Despite incurring a significant degradation in axial resolution, the image of Fig. 9(b) still shows visible sidelobe artefacts demonstrating the limitations of windowing as a sidelobe suppression technique.



**Fig. 9.** Comparison of spectral windowing with CSC for sidelobe artefact suppression for the index finger (palmar aspect) image. Windows are applied on the full 1460 sample spectrograms then truncated to 330 pixels following inverse discrete Fourier transformation. Separately, a mirror reflector was imaged to provide direct measurements of the PSF and corresponding spectrogram. The insets show pixel intensities in regions  $H_2$  and  $A$ . (a) image with no spectral windowing applied. (b) Image with Gaussian windowing ( $\sigma = 0.10$  of spectrogram width). (c) Image with Hann windowing. (d) proposed convolutional sparse coding method with a PSF learned from the image. (e) Measured PSF from a mirror reflector with no window, Hann window and Gaussian window ( $\sigma = 0.10$  of spectrogram width) applied. The arrows indicate features that contribute to the sidelobe artefacts and that are not suppressed by windowing. (f) Spectrogram with no window, Hann window and Gaussian windows applied. The plotted spectrogram is the average over 15,000 laser sweeps. The PSFs in (e) are obtained by inverse Fourier transformation of the spectrograms in (f).

## 6. Conclusion

Blind deconvolution using convolutional sparse coding with dictionary learning is a promising approach for reducing artefacts and noise and for enhancing contrast in OCT images containing bright reflectors where contrast is limited by sidelobe artefacts from the PSF. For a fingertip image, we demonstrated an SNR improvement of 2.2 dB and a contrast improvement of 5.8 dB within a region affected by PSF sidelobes while only incurring a 0.3 dB contrast change between a brightly reflecting and a weakly reflecting tissue region. The proposed processing also improved the generalized contrast-to-noise ratio between tissue and background, and between tissue and artefact regions, but reduced gCNR between the bright and weak tissue regions due to the introduction of zero-intensity pixels into the speckle in the bright tissue region. We showed that this effect can be largely mitigated by applying a despeckling step (e.g., median filtering) after deconvolution.

We believe this to be the first demonstration of blind deconvolution of complex OCT A-lines. Unlike previously described deconvolution methods for OCT [26–28], the improvements in image quality we obtain do not require us to obtain an independent measurement or model-based estimate of the PSF. Rather, a complex valued, axial PSF is learned from the complex OCT A-lines acquired during imaging. This makes the method particularly well-suited to OCT systems in which the PSF changes with time because dictionary learning can continuously update its estimate of the PSF from OCT imaging data without requiring any additional calibration steps.



The stability of the PSF will depend on the details of the source and interferometer design of each OCT system. In our swept-source system we found that the PSF estimate could be used to effectively suppress sidelobe artefacts over a period of tens of seconds before requiring updating. It also makes the method applicable to previously collected image data, so long as the complex line data or the spectrogram data was retained.

Our current CSC processing rate of 2.5 ms/line (1.28s for a 512-line image) is suitable for many near-real-time image post-processing applications such as post-rendering for 3D visualization, image segmentation and feature extraction. Further improvements to speed using GPU accelerated algorithms could plausibly make the method compatible with real-time, video-rate 2D processing for OCT images of typical size. While CDL is a slower step (18s for 128 lines) than CSC, it only needs to be performed periodically and so it may also be compatible with real-time processing rates with GPU acceleration. For real-time processing, a modified form of the CDL problem that performs continuous and efficient updating of the PSF estimate called online CDL may be applied [70].

Effective use of convolutional sparse coding requires the appropriate selection of the regularization parameter  $\lambda$  and  $\ell_1$  weighting factor  $W$ . While this optimization was done heuristically in this study, performance was only weakly dependent on the exact values chosen and values of  $\lambda$  in the range of 0.01 to 0.05 and  $W = 0.1$  worked well across the various images we tested.

The removal of sidelobe artefacts and noise from OCT images containing bright reflectors is an important step for allowing advanced visualization and analysis steps including 3D volumetric visualization, automated segmentation, and automated diagnostics. Blind deconvolution through convolutional sparse coding with dictionary learning could prove to be an important image processing technique for many OCT applications.

**Funding.** Los Alamos National Laboratory (20200061DR); Natural Sciences and Engineering Research Council of Canada (2020-06321); Mitacs (IT14716).

**Acknowledgements.** The authors gratefully acknowledge support for this project from Michelle Bona, Josh Farrell, Matthew Farrell, Alex Hackett, Drew Hubley, Matthew Jahns, Kaila Kelly, Dan MacDougall, Darren Oickle.

**Disclosures.** J.W: Audioprocs Medical Inc. (F, I, E), RBA: Audioprocs Medical Inc. (F, I, E)  
B. Wohlberg declares that there are no conflicts of interest related to this article.

**Data availability.** Data and computer source code used to generate the results presented in this paper are available in Ref. [60].

## References

1. W. Drexler and J. G. Fujimoto, *Optical Coherence Tomography: Technology and Applications* (Springer International Publishing, 2015).
2. X. Yu, X. Liu, J. Gu, D. Cui, J. Wu, and L. Liu, "Depth extension and sidelobe suppression in optical coherence tomography using pupil filters," *Opt. Express* **22**(22), 26956 (2014).
3. E. Bousi and C. Pitris, "Axial resolution improvement by modulated deconvolution in Fourier domain optical coherence tomography," *J. Biomed. Opt.* **17**(7), 071307 (2012).
4. Y. Wang, Y. Liang, and K. Xu, "Signal processing for sidelobe suppression in optical coherence tomography images," *JOSA A* **27**(3), 415 (2010).
5. C. Pitris, E. P. Ippen, F. X. Kärtner, J. G. Fujimoto, S. A. Boppart, U. Morgner, W. Drexler, and X. D. Li, "In vivo ultrahigh-resolution optical coherence tomography," *Opt. Lett.* **24**(17), 1221–1223 (1999).
6. M. Bonesi, M. P. Minneman, J. Ensher, B. Zabihian, H. Sattmann, P. Boschert, E. Hoover, R. A. Leitgeb, M. Crawford, and W. Drexler, "Akinetic all-semiconductor programmable swept-source at 1550 nm and 1310 nm with centimeters coherence length," *Opt. Express* **22**(3), 2632 (2014).
7. A. F. Fercher, W. Drexler, C. K. Hitzenberger, and T. Lasser, "Optical coherence tomography - principles and applications," *Rep. Prog. Phys.* **66**(2), 239–303 (2003).
8. B. C. Quirk, R. A. McLaughlin, A. Curatolo, R. W. Kirk, P. B. Noble, and D. D. Sampson, "In situ imaging of lung alveoli with an optical coherence tomography needle probe," *J. Biomed. Opt.* **16**(3), 036009 (2011).
9. Z. Hamdoun, W. Jerjes, R. Al-Delayme, G. McKenzie, A. Jay, and C. Hopper, "Structural validation of oral mucosal tissue using optical coherence tomography," *Head Neck Oncol.* **4**(1), 29 (2012).
10. J. A. Burns, "Optical coherence tomography: imaging the larynx," *Curr. Opin. Otolaryngol. Head Neck Surg.* **20**(6), 477–481 (2012).



11. M. Machoy, J. Seeliger, L. Szyszka-Sommerfeld, R. Koproński, T. Gedrange, and K. Woźniak, "The use of optical coherence tomography in dental diagnostics: A state-of-the-art review," *J. Healthc. Eng.* **2017**, 1–31 (2017).
12. D. MacDougall, J. Farrell, J. Brown, M. Bance, and R. Adamson, "Long-range, wide-field swept-source optical coherence tomography with GPU accelerated digital lock-in Doppler vibrography for real-time, in vivo middle ear diagnostics," *Biomed. Opt. Express* **7**(11), 4621 (2016).
13. D. MacDougall, J. Rainsbury, J. Brown, M. Bance, and R. Adamson, "Optical coherence tomography system requirements for clinical diagnostic middle ear imaging," *J. Biomed. Opt.* **20**(5), 056008 (2015).
14. Thorlabs Inc, "OCT Selection Guide," [https://www.thorlabs.com/newgrouppage9.cfm?objectgroup\\_id=5702](https://www.thorlabs.com/newgrouppage9.cfm?objectgroup_id=5702).
15. Wasatch Photonics, "OCT Image | OCT Angiography | Gallery," <https://wasatchphotonics.com/oct-image-gallery/>.
16. W. Kim, S. Kim, S. Huang, J. S. Oghalai, and B. E. Applegate, "Picometer scale vibrometry in the human middle ear using a surgical microscope based optical coherence tomography and vibrometry system," *Biomed. Opt. Express* **10**(9), 4395 (2019).
17. Z. Hamdoon, W. Jerjes, G. McKenzie, A. Jay, and C. Hopper, "Optical coherence tomography in the assessment of oral squamous cell carcinoma resection margins," *Photodiagn. Photodyn. Ther.* **13**, 211–217 (2016).
18. K. Al-Azri, L. N. Melita, A. P. Strange, F. Festy, M. Al-Jawad, R. Cook, S. Parekh, and L. Bozec, "Optical coherence tomography use in the diagnosis of enamel defects," *J. Biomed. Opt.* **21**(3), 036004 (2016).
19. E. Sattler, R. Kästle, and J. Welzel, "Optical coherence tomography in dermatology," *J. Biomed. Opt.* **18**(6), 061224 (2013).
20. M. Choma, M. Sarunic, C. Yang, and J. Izatt, "Sensitivity advantage of swept source and Fourier domain optical coherence tomography," *Opt. Express* **11**(18), 2183 (2003).
21. J. Kalkman, "Fourier-domain optical coherence tomography signal analysis and numerical modeling," *Int. J. Opt.* **2017**, 1–16 (2017).
22. D. P. Popescu, L.-P. Choo-Smith, C. Flueraru, Y. Mao, S. Chang, J. Disano, S. Sherif, and M. G. Sowa, "Optical coherence tomography: Fundamental principles, instrumental designs and biomedical applications," *Biophys. Rev.* **3**(3), 155–169 (2011).
23. P. H. Tomlins and R. K. Wang, "Theory, developments and applications of optical coherence tomography," *J. Phys. D: Appl. Phys.* **38**(15), 2519–2535 (2005).
24. D. Piao, Q. Zhu, N. K. Dutta, S. Yan, and L. L. Otis, "Cancellation of coherent artifacts in optical coherence tomography imaging," *Appl. Opt.* **40**(28), 5124 (2001).
25. Y. Chen, J. Fingler, and S. E. Fraser, "Multi-shaping technique reduces sidelobe magnitude in optical coherence tomography," *Biomed. Opt. Express* **8**(11), 5267 (2017).
26. S. A. Hojjatoleslami, M. R. N. Avanaki, and A. G. Podoleanu, "Image quality improvement in optical coherence tomography using Lucy–Richardson deconvolution algorithm," *Appl. Opt.* **52**(23), 5663 (2013).
27. P. D. Woolliams, R. A. Ferguson, C. Hart, A. Grimwood, and P. H. Tomlins, "Spatially deconvolved optical coherence tomography," *Appl. Opt.* **49**(11), 2014 (2010).
28. J. M. Schmitt, "Restoration of optical coherence images of living tissue using the CLEAN algorithm," *J. Biomed. Opt.* **3**(1), 66 (1998).
29. G. Liu, S. Yousefi, Z. Zhi, and R. K. Wang, "Automatic estimation of point-spread-function for deconvoluting out-of-focus optical coherence tomographic images using information entropy-based approach," *Opt. Express* **19**(19), 18135 (2011).
30. D. Kundur and D. Hatzinakos, "Blind image deconvolution," *IEEE Signal Process. Mag.* **13**(3), 43–64 (1996).
31. A. A. Moiseev, G. V. Gelikonov, P. A. Shilyagin, and V. M. Gelikonov, "Blind deconvolution algorithm for restoration OCT images with diffraction limited resolution," in R. A. Leitgeb and B. E. Bouma, eds. (SPIE, 2011), 8091, p. 80911W.
32. K. Shen, H. Lu, S. Baig, and M. R. Wang, "Improving lateral resolution and image quality of optical coherence tomography by the multi-frame superresolution technique for 3D tissue imaging," *Biomed. Opt. Express* **8**(11), 4887 (2017).
33. B. Wohlberg and P. Wozniak, "PSF Estimation in crowded astronomical imagery as a convolutional dictionary learning problem," *IEEE Signal Processing Letters* **28**, 374–378 (2021).
34. D. L. Donoho and M. Elad, "Optimally sparse representation in general (nonorthogonal) dictionaries via  $L_1$  minimization," *Proc. Natl. Acad. Sci.* **100**(5), 2197–2202 (2003).
35. L. Fang, S. Li, Q. Nie, J. A. Izatt, C. A. Toth, and S. Farsiu, "Sparsity based denoising of spectral domain optical coherence tomography images," *Biomed. Opt. Express* **3**(5), 927 (2012).
36. Y. C. Eldar, *Compressed Sensing* (Cambridge University, 2012).
37. S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proc. Natl. Acad. Sci.* **113**(15), 3932–3937 (2016).
38. B. Wohlberg, "Efficient algorithms for convolutional sparse representations," *IEEE Trans. Image Process.* **25**(1), 301–315 (2016).
39. Z. Zhang, Y. Xu, J. Yang, X. Li, and D. Zhang, "A survey of sparse representation: algorithms and applications," *IEEE Access* **3**, 490–530 (2015).
40. B. A. Olshausen, "Learning sparse, overcomplete representations of time-varying natural images," in *Proceedings 2003 International Conference on Image Processing* (IEEE, 2003), 1, p. I-41–4.

41. W. Hashimoto and K. Kurata, "Properties of basis functions generated by shift invariant sparse representations of natural images," *Biol. Cybern.* **83**(2), 111–118 (2000).
42. M. Lewicki and T. Sejnowski, "Coding time-varying signals using sparse, shift-invariant representations," in *Advances in Neural Information Processing Systems 11 (NIPS 1998)* (1998).
43. A. Cogliati, Z. Duan, and B. Wohlberg, "Context-dependent piano music transcription with convolutional sparse coding," *IEEE/ACM Transactions on Audio Speech, and Language Processing* **24**(12), 2218–2230 (2016).
44. A. Cogliati, Z. Duan, and B. Wohlberg, "Piano music transcription with fast convolutional sparse coding," in *2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)* (IEEE, 2015), 2015–Novem, pp. 1–6.
45. N. Mohan, I. Stojanovic, W. C. Karl, B. E. A. Saleh, and M. C. Teich, "Compressed sensing in optical coherence tomography," in *Three-Dimensional and Multidimensional Microscopy: Image Acquisition and Processing XVII*, J.-A. Conchello, C. J. Cogswell, T. Wilson, and T. G. Brown, eds. (2010), 7570, p. 75700L.
46. C. Liu, A. Wong, K. Bizheva, P. Fieguth, and H. Bie, "Homotopic, non-local sparse reconstruction of optical coherence tomography imagery," *Opt. Express* **20**(9), 10200 (2012).
47. M. Shamouilian and I. Selesnick, "Total variation denoising for optical coherence tomography," in *2019 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)* (IEEE, 2019), pp. 1–5.
48. L. Fang, S. Li, X. Kang, J. A. Izatt, and S. Farsiu, "3-D adaptive sparsity based image compression with applications to optical coherence tomography," *IEEE Trans. Med. Imaging* **34**(6), 1306–1320 (2015).
49. Y. Ling, M. Wang, Y. Gan, X. Yao, L. Schmetterer, C. Zhou, and Y. Su, "Beyond Fourier transform: super-resolving optical coherence tomography," (2020).
50. L. Vinet and A. Zhedanov, "A 'missing' family of classical orthogonal polynomials," *J. Phys. A: Math. Theor.* **44**(8), 085201 (2011).
51. B. Wohlberg, "Efficient convolutional sparse coding," in *2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (IEEE, 2014), pp. 7173–7177.
52. H. Hongxing, J. M. Bioucas-Dias, and V. Katkovnik, "Interferometric phase image estimation via sparse coding in the complex domain," *IEEE Trans. Geosci. Remote Sens.* **53**(5), 2587–2602 (2015).
53. X. Liu and J. U. Kang, "Compressive SD-OCT: the application of compressed sensing in spectral domain optical coherence tomography," *Opt. Express* **18**(21), 22010 (2010).
54. A. Cameron, D. Lui, A. Boroomand, J. Glaister, A. Wong, and K. Bizheva, "Stochastic speckle noise compensation in optical coherence tomography using non-stationary spline-based speckle noise modelling," *Biomed. Opt. Express* **4**(9), 1769 (2013).
55. C. Garcia-Cardona and B. Wohlberg, "Convolutional dictionary learning: A comparative review and new algorithms," *IEEE Trans. Comput.* **4**(3), 366–381 (2018).
56. C. Garcia-Cardona and B. Wohlberg, "Subproblem coupling in convolutional dictionary learning," in *2017 IEEE International Conference on Image Processing (ICIP)* (IEEE, 2017), (3), pp. 1697–1701.
57. J. F. C. Mota, J. M. F. Xavier, P. M. Q. Aguiar, and M. Puschel, "Distributed basis pursuit," *IEEE Trans. Signal Process.* **60**(4), 1942–1956 (2012).
58. J. Kang, D. Hong, J. Liu, G. Baier, N. Yokoya, and B. Demir, "Learning convolutional sparse coding on complex domain for interferometric phase restoration," *IEEE Trans. Neural Netw. Learn. Syst.* **32**(2), 826–840 (2021).
59. B. Wohlberg, "SPORCO: A Python package for standard and convolutional sparse representations," in *Proceedings of the 16th Python in Science Conference (SciPy, 2017)*, pp. 1–8.
60. J. Wang and R. Adamson, "young-oct/OCT-sparse-estimation-with-CBPDN-framework," Github, 2022, <https://github.com/young-oct/OCT-sparse-estimation-with-CBPDN-framework>.
61. S. Moon, S.-W. Lee, and Z. Chen, "Reference spectrum extraction and fixed-pattern noise removal in optical coherence tomography," *Opt. Express* **18**(24), 24395 (2010).
62. G.-M. Zhang, C.-Z. Zhang, and D. M. Harvey, "Sparse signal representation and its applications in ultrasonic NDE," *Ultrasonics* **52**(3), 351–363 (2012).
63. R. C. Gonzalez and R. E. Woods, "Digital Image Processing, Global Edition," in *Digital Image Processing*, Global Edition (2017), 19(3), p. 1024.
64. Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Trans. on Image Process.* **13**(4), 600–612 (2004).
65. Han Chumming, Guo Huadong, and Wang Changlin, "Edge preservation evaluation of digital speckle filters," in *IEEE International Geoscience and Remote Sensing Symposium* (IEEE, 2002), 4, pp. 2471–2473.
66. C. H. Gierull and I. C. Sikaneta, "Estimating the effective number of looks in interferometric SAR data," *IEEE Trans. Geosci. Remote Sens.* **40**(8), 1733–1742 (2002).
67. K. Wang, Z. Ding, M. Chen, C. Wang, T. Wu, and J. Meng, "Deconvolution with fall-off compensated axial point spread function in spectral domain optical coherence tomography," *Opt. Commun.* **284**(12), 3173–3180 (2011).
68. K. M. Kempfski, M. T. Graham, M. R. Gubbi, T. Palmer, and M. A. Lediju Bell, "Application of the generalized contrast-to-noise ratio to assess photoacoustic image quality," *Biomed. Opt. Express* **11**(7), 3684 (2020).

69. A. Rodriguez-Molares, O. M. H. Rindal, J. D'hooge, S.-E. Masoy, A. Austeng, M. A. Lediju Bell, and H. Torp, "The generalized contrast-to-noise ratio: a formal definition for lesion detectability," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control* **67**(4), 745–759 (2020).
70. J. Liu, C. Garcia-Cardona, B. Wohlberg, and W. Yin, "Online convolutional dictionary learning," in *2017 IEEE International Conference on Image Processing (ICIP)* (IEEE, 2017), 2017-Sep(4), pp. 1707–1711.